DOCUMENT RESUME

ED 032 778

EM 007 450

By-Stolurow, Lawrence M.; Klare, George R.

Project CREATES. First Annual Report 1968-69.

Harvard Univ., Cambridge, Mass. Computation Lab.
Spons Agency-Office of Education (DHEW), Washington, D.C.
Report No-ESEA-3; PL-89-750
Pub Date 15 Jun 69
Grant-OEG-0-8-031083-4369-039
Note-59p.; Appendix not available from EDRS
Available from-(including appendix) ERIC Clearinghouse on Adult Education, Syracuse University, 107 Roney Lane, Syracuse, New York 13210
EDRS Price MF-$0.50 HC-$3.05
Descriptors-*Adult Basic Education, Adult Education, Adult Education Programs, Adult Learning, *Adult Reading Programs, Cloze Procedure, Computer Assisted Instruction, Computer Oriented Programs, Curriculum Development, *Curriculum Evaluation, *Instructional Materials, Instructional Materials Centers, Language Instruction, Publications, Readability, *Reading Instruction, Reading Material Selection, Reading Research
Identifiers-Combined Resources for Editing Automated Teaching, CREATES, Harvard University Computer Aided Instruction

Project CREATES stands for Combined Resources for Editing Automated Teaching to Enlighten Students, the Harvard University Computer-Aided Instruction Laboratory's part of "A Multi-Agency Developmental Adult Basic Education Program." The major emphases of Project CREATES have been in the area of program and procedure development for computer-aided revision of instructional materials and the area of computer instruction of reading and language skills. Work has also been done on analytic-generative computer programs to aid editors and writers in text revision, and on the development of new programs for Adult Basic Education. The Computer Aided Instruction training reported was carried out in the Learning Center established in the Roxbury section of Boston. The first training session was still in process when the report was written, and definite test results were unobtainable. However, the work done by Project CREATES personnel from the inception of the program to the time of writing are covered in detail under the categories: Personnel, Publications, Meetings, and Research and Development. An appendix containing most of the publications, technical reports and notes mentioned in the report is obtainable from Syracuse University. A glossary of acronyms, Personnel list, and a schematic diagram of the program are included. (SH)

ED0 32778

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

First Annual Report (1968-69).

Project CREATES

Title III USEA Contract No. OEG-0-8-031083-4369(039)

Lawrence M. Stolurow and George R. Klare

June 15, 1969

EM 007 450

Principal Investigator:
Lawrence M. Stolurow

46 Concord Avenue
Cambridge, Mass. 02138

Table of Contents

List of Figures.....i

Introduction.....1

Purpose.....1

Locations.....3

Glossary.....4

Organization of Report.....5

Personnel.....6

Publications.....7

Journal and Newsletter Publications.....7

Technical Reports.....7

Technical Notes.....8

Progress Reports.....8

Training Memoranda and Forms.....9

Miscellaneous Forms and Diagrams.....10

Meetings.....11

Research and Development.....13

Sub-Project TREE.....15

Methods for Evaluating Changes.....15

Word Changes.....19

Sentence Changes.....22

Other Changes.....25

Instructional Inputs.....27

Rewriting Systems (automation levels).....32

Instructional Outputs (applications).....36

Sub-Project CAIBLS.....39

Materials Development Stage.....40

Table of Contents (cont.)

Student Acceptance Stage.....46

Staff Acceptance Stage.....47

Curriculum Stage.....48

Generalization Stage.....50

Appendix.....55

List of Figures

Figure 1, An Organization Scheme for Project CREATES...14

Figure 2, Evaluation Form, CARET I System.....34-35

Figure 3, Lessons on the CAI System.....40-41

Project CREATES

Adult Basic Education

Annual Report of September 1, 1968 - June 15, 1969, Contract Period

L. M. Stolurow and George R. Klare

Introduction

This report covers work on Project CREATES during the first contract period, September 1, 1968 - June 15, 1969. Project CREATES is Harvard University Computer-Aided Instruction Laboratory's part of "A Multi-Agency Developmental Adult Basic Education Program." This is a special experimental demonstration project supported by a grant from the U. S. Office of Education under the Adult Education Act of 1966, P. L. 89-750 (Special Project Grant Title III, P. L. 89-750). The grant was made to the Massachusetts Department of Education, with Bruno Ciccariello serving as overall Project Director. Lawrence M. Stolurow served as Principal Investigator and George R. Klare served as Project Director for Harvard University.

Purpose

Project CREATES consists of two closely-related parts, Sub-Project TREE and Sub-Project CAIBLS. The designation TREE comes from the phrase "Translation from English to English," a shorthand term for the first overall project emphasis upon development of programs and procedures for computer-aided revision of instructional materials to meet the limited capabilities of adult basic education (ABE) students. The designation CAIBLS comes from the phrase "Computer-Aided Instruction in Basic Language Skills," a shorthand term for the second overall project emphasis upon use of the computer as an

instructional device in reading and related language instruction. TREE and CAIBLS are most obviously related in that CAI instructional materials are to be rewritten through programs developed under the former sub-project for use in the latter sub-project.

In addition, several newer emphases have heightened the relationship. In order to develop, evaluate, and refine the revision programs, it is necessary to develop certain "analytic" programs to provide an editor or writer with information about text that is to be revised. It was found, as work progressed, that properly developed analytic programs could often also serve a "generative" function. That is, instructional materials could themselves be generated by such programs, and could thus feed into the computer-aided instruction (CAIBLS) part of the project. The first of the newer emphases, then, was work on analytic-generative computer programs. The work has been planned under the general heading (for the entire system) of PAGES (Psychological-Analytic-Generative-Educational System). This system will eventually embrace the many and varied programs being developed under the Project.

A second emphasis was on new programs for which the ABE Learning Center staff expressed a need after their work had begun. Among these needs was that for mathematics programs, to which the CAI Lab responded with programs covering basic arithmetic operations and the understanding of fractions. Another need was for programs in certain basic language skills to accompany the "story" lessons developed to teach vocabulary, comprehension, following of directions, etc. Programs in the teaching of phonic skills and spelling are being developed to meet these needs. These newer programs have

additional characteristics discovered, during the CAI work with the ABE students, to be of importance for most effective instruction.

Briefly, then, the emphases of the CREATES project have been the following.

(1) Sub-project TREE - development of a computer-aided system for revision of instructional materials to meet the capabilities of ABE students; also, development of related analytic programs for text analysis, and for computer-aided generation of new instructional materials.

(2) Sub-project CAIBLS - development of basic language materials for computer-aided instruction; also, development of special CAI skills programs in both language and mathematics designed especially to meet the demonstrated needs of ABE students.

Locations

The CAI training has been carried out in the Learning Center established at the Opportunities Industrialization Center (OIC) on Dudley Street in the Roxbury section of Boston. Adult basic education students have used the CAI console since January, 1969, as part of their training. The first session of this training will have ended on approximately July 3 (the Project has been extended from June 16, 1969 to June 30, 1970). Certain data analyses will be made after the last part of the training has been completed; since this will be in July, they could not be completed at the time this report was written, but will be included in a subsequent report.

Note that plans are being completed at this time, also, for an additional Learning Center in Boston to be operated by the Adult

Basic Education Department of the Boston Public Schools (as part of a program for the Work Incentive - WIN - Program of the U. S. Department of Labor). This new Center, and perhaps others, will join the OIC Learning Center during the second year's work.

Glossary

It may be desirable, at this point, to present a brief glossary of the various abbreviations (mostly acronyms) already used in the report as well as others still to come. Those who work with computers find that meaningful abbreviations of this sort provide an efficient reference method (as well as a challenge to their creative ingenuity). If the acronyms sometimes prove more puzzling than helpful to readers, perhaps this glossary will provide a useful reference.

1. ABE — Adult Basic Education
2. ALAN — Adult LANguage (a set of adult instructional programs)
3. BACON — BAsic CONcordance (a concordance program)
4. CAI — Computer Aided (or Assisted) Instruction (may appear in CAI Lab, referring to the Harvard Computer-Aided Instruction Laboratory)
5. CAIBLS — Computer Aided Instruction in Basic Language Skills (a sub-project of CREATES)
6. CLIP — CLoze Instructional Program
7. CREATES — Combined Resources for Editing Automated Teaching to Enlighten Students
8. FRECAI — Flesch Reading Ease, Computer Aided Instruction (a version of a general program)
9. HALF — Refers to an instructional program to teach fractions
10. KWIC — Key-Word-In-Context (a concordance program)
11. OIC — Opportunities Industrialization Center
12. PAGES — Psychological-Analytic-Generative-Educational System (a set of programs)

13. SOCCER — Smart's Own ConCordance, Extremely Rapid (a concordance program)

14. TREE — TRanslation from English to English (a sub-project of CREATES)

15. URLAN — URban LANguage (a set of adult instructional programs)

16. WIN — Work INcentive project (an adult education project sponsored by the Department of Labor)

17. WISSYN — WISconsin SYNonym program (a surface parsing program)

Organization of Report

The remainder of this report will present work done by Project CREATES personnel during the September 1, 1968 - June 15, 1969 period. The report will be organized under the following categories: Personnel; Publications; Meetings; and Research and Development. Materials cited in the report will be included in an appendix.

Personnel

The following persons worked full-time or part-time on Project CREATES during all or part of the period September 1, 1968 - June 15, 1969. Their titles indicate, insofar as possible, the nature of their contribution to the Project. Several of the program analysts were graduate students who were doing this work as part of their graduate training.

(1) Lawrence M. Stolurow - Principal Investigator

(2) George R. Klare - Project Director

(3) Blanche Serwer - Language Arts Consultant

(4) Theodore Peterson - IBM Fellow

(5) M. Gregory St. John - Computer Programmer

(6) Paul P. Rowe - Computer Programmer

(7) Georgina von Hindy - Language Arts Specialist

(8) Walter Daugherity - Systems Analyst

(9) Thomas Manwell - Systems Analyst

(10) Louise Oremland - Language Program Analyst

(11) Sharon Curry - Language Program Analyst

(12) Alan Tung - Mathematics Program Analyst

(13) Marcia Bandas - Language Program Analyst

(14) Kathleen Gerritz - Language Program Analyst

(15) Theresa Auerbach - Language Program Analyst

(16) Lucy Carroll - Language Program Analyst

(17) Leonard Glick - Mathematics Program Analyst

Secretarial and related assistance came from Elizabeth Walker, Greta Wiseblatt, Arden Howard, Sally Winters, Margaret Stevens, and others.

Publications

This heading is meant to include the following categories of items: journal and newsletter publications; technical reports; technical notes; progress reports; training memoranda questionnaires and forms; and miscellaneous forms and diagrams. Items will simply be listed here; the research and development work reported in them will be summarized under the Research and Development heading below. Copies of each of the above can be found in the appendix to this report, or can be obtained from the ERIC Clearinghouse.

Journal and Newsletter Publications

These are reports in printed form intended to reach an audience of professional readers in either a research area (i.e., readers of a technical journal) or a teaching area (i.e., adult basic education personnel).

(1) Automation of the Flesch "Reading Ease" Readability Formula, with Various Options (G. R. Klare, P. P. Rowe, M. G. St. John, and L. M. Stolurow). Reading Research Quarterly, Vol. IV, No. 4, Summer, 1969, pages 550 - 559.

(2) A Computer-Aided Instructional Approach to Support Adult Learning Centers (G. R. Klare). A. B. E. Newsletter, Mass. Dept. of Education and U. S. Office of Education, February, 1969. (Appeared previously in memorandum form as: An Instructional Approach to Support Adult Learning Centers.)

Technical Reports

These are research and development documents which describe major outcomes of Project work.

(1) Technical Report No. I.1.1: Automation of the Flesch "Reading Ease" Readability Formula, CAI Version (FRECAI), January 20,

1969 (G. R. Klare and M. G. St. John).

(2) Technical Report No. B.1.1: Words, Sentences, and Readability, May 5, 1969 (G. R. Klare).

(3) Technical Report No. B.2.1: A First Version of a Computer-Aided Revising, Editing, and Translating System (CARET I), June 15, 1969 (G. R. Klare, P. P. Rowe, M. G. St. John, and L. M. Stolurow).

Technical Notes

These are research and development documents which describe preliminary outcomes of Project work.

(1) Technical Note No. 1: An Example of Proposed Text Analysis, October 1, 1968 (G. R. Klare).

(2) Technical Note No. 2: Descriptive and Evaluative Materials Prepared for Instructors at the OIC Learning Center, January 15, 1969 (G. R. Klare).

(3) Technical Note No. 3: A Preliminary Study of the Reading Interests of OIC Learning Center Students, March 29, 1969 (G. R. Klare).

Progress Reports

These are reports prepared monthly to provide information on Project work done and next steps planned. Included also are "Explanatory Notes" designed to cover particular points or projects to which attention is to be called. Since Progress Report No. 1 and Progress Report No. 1 (Revised) contain a diagram around which all the remaining work and reports are organized, one or the other should be examined before the later reports are read.

(1) Progress Report No. 1 - October 21, 1968.

(2) Progress Report No. 1 (Revised) - November 1, 1968.

(3) Progress Report No. 2 - December 1, 1968.

(4) Progress Report No. 3 - January 2, 1969.

(5) Progress Report No. 4 - February 1, 1969.

(6) Progress Report No. 5 - March 1, 1969.

(7) Progress Report No. 6 - April 1, 1969.

(8) Progress Report No. 7 - May 1, 1969.

(9) Progress Report No. 8 - June 1, 1969.

Training Memoranda and Forms

These are materials intended for use in the pre-service training of personnel who are to operate the CAI console, or to know of its operation so that they can assign students to it.

(1) Memorandum dated 1/2/69, "Use of Lessons on Computer Terminal (Console)."

(2) Memorandum dated 1/13/69, "Lessons Available for Use at OIC."

(3) Memorandum (no date), "Procedure for Signing On."

(4) Memorandum dated 12/11/68, "Trouble Check-List, Draft 0."

(5) Memorandum (no date), "Problems You Can Avoid (And How)."

(6) Memorandum (no date), "How to Monitor vincon (For Example)."

(7) Memorandum dated 1/3/69, "The RPTest."

(8) Form dated 12/31/68, "Student Record Form."

(9) Form dated 12/17/68, "Scale of Attitudes Toward Computers for Boston Students," version PRE.

(10) Form dated 12/17/68, "Scale of Attitudes Toward Computers for Boston Students," version POST.

(11) Form (no date), "Teacher Questionnaire for CAI Unit Lesson."

(12) Form dated 1/2/69, "Reading Interest Questionnaire," for use with students or teachers.

(13) List (no date), "Descriptions of Stories." This list

contains basically the same information as the form in number 12 (above), so only one or the other is needed.

(14) List (no date), "URLAN Course List."

(15) List (no date or title) of all stories and tests prepared or being prepared for computer use. This list contains much the same information as the list in number 14 (above), so only one or the other is needed.

(16) Memorandum (no date) covering "Half," a program to teach fractions.

(17) Memorandum (no date) covering "URAE02," a program to teach basic arithmetic operations.

Miscellaneous Forms and Diagrams

(1) Some Instructional Uses of Computers (diagram) (no date).

(2) Examples of Some Current Program Output (diagram) (no date).

(3) Reading Material Matrix form (no date).

(4) Project CREATES - three diagrams showing: Computers in Education; general project organization; general project procedures.

Meetings

Since a very large number of meetings were held, no attempt will be made to note every one. Instead, major meetings and categories of meetings will be listed.

(1) Preliminary meetings with OIC and Board of Education personnel. A number of such meetings were held for acquaintance and then for general planning purposes. Early meetings provided a work schedule to be followed. Later meetings brought certain OIC supervisors, teachers, and students to the CAI Lab to try the console before the Learning Center program was started.

(2) Pre-Service Training meetings. The first of such meetings were devoted to planning the Pre-Service Training Program. Later meetings over a 2-week period were devoted to the Program itself. The Harvard portion of the latter were held at the CAI Lab so that a console would be available.

(3) Communication meetings. Approximately 8 such meetings were held so that the various contributors to the total Project could be informed of each others' work.

(4) Planning meetings for other Learning Centers. Approximately 6 meetings were held to discuss other locations for Learning Centers.

(5) Meetings with technical specialists. Approximately 12 meetings were held with specialists who could contribute knowledge to the project operations.

(6) Guest visits. A number of guests, many foreign, visited the CAI Laboratory in order to observe the use of the console and the overall instructional work.

(7) Evaluation visits. A major evaluation visit was paid by

Mr. Gene Sullivan and Mr. Bayard Clark of the U. S. Office of Education on June 5 - 6, 1969. Several visits were also carried out by members of the Heuristics organization, which is providing an on-going analysis and evaluation of the overall Project.

(8) Discussion and professional meetings. The work of the CAI Laboratory on Project CREATES was discussed in various meetings, such as a group of principals meeting at OIC, a group of adult educators meeting in Lexington, Massachusetts, and another group of educators meeting at the Harvard Faculty Club. Papers were presented by five persons on Project work at the New England Educational Research Organization meeting held at Boston College on June 5 - 6, 1969. Two additional meetings were planned for July, 1969, and another for December, 1969.

Research and Development

This section will present a summary of the main body of work done during the period September 1, 1968 - June 15, 1969. It has been organized around the plan developed for the first and subsequent progress reports. Each of the headings will be described here, as an expansion of the brief headings used in the figure and elsewhere in the first progress report. For the convenience of the reader, the figure is reproduced here as Figure 1, An Organizational Scheme for Project CREATES Work. The rationale for Figure 1 is that the four diagonal legs, numbered I to IV, represent major directions of research and development work. The three perpendicular legs lettered A to C represent the TREE process, from instructional input through revision, to instructional output. (The fourth of the perpendicular legs, unlettered, serves simply for designation purposes.) The goal of the project as a whole, as indicated by the central rectangle, is "Changes in Instructional Efficiency." The authors hope that the various headings are themselves descriptive enough to stand alone for the reader who prefers not to follow through the figure itself.

The first set of 7 major headings below (I to IV and A to C) will refer primarily to Sub-project TREE work and the second set of 5 (I to V) to Sub-project CAIBLS work. The organizational scheme followed for this latter part of the report was presented as the Explanatory Note section of Progress Report No. 3, January 2, 1968. Note that, as mentioned above, data analyses for CAIBLS will not be included here owing to the fact the first group will not finish its training until early July. Such analyses will therefore be included in a later report.

I. Methods for Evaluating Changes:

1. Readability Measurement (Flesch&Dale-Chall Formulas)

2. Cloze Measurement

3. Speed of Reading Measurement

4. Acceptibility Measurement (Student & Teachers)

A. Instructional Inputs: (Reading Materials Corpus)

1. Grade Level by Interest Corpus

2. Semi-technical Corpus

3. General Inquirer Corpus

II. Word Changes:

1. General Inquirer & SOCCER Program Trials

2. KWIC and Frequency Listings

3. Sense Rules

4. Word Length and Frequency Changes

Harvard CAI Lab — Project CREATES — Project TREE, Project CAIBLS

Changes in Instructional Efficiency

2. Computer Assisted Revision — 1. Computer Assisted Explanation — B. Rewriting Systems (Automation Levels)

4. Transformation Difficulty

3. Phrase & Clause Pattern Difficulty

2. Technical Term Definitions

1. Nominalization Changes

IV. Other Changes:

3. Rewritten Materials for Training Use

2. New Instructional Techniques

1. Evaluation of Existing Adult Materials

C. Instructional Outputs: (Applications)

4. Sentence Length and Complexity Changes

3. Pronoun Changes

2. Form-class Listings

1. WISSYN Program Trial

III. Sentence Changes

Sub-Project TREE

I. Methods for Evaluating Changes. This heading refers primarily to various methods or procedures to be used in evaluating the effectiveness of changes made in instructional programs for greater teaching efficiency. The work done in each area will be described below.

(1) Readability measurement. Two wide-range readability analysis programs were planned, based on a review of the readability literature. It was found that if analyses of three basic language variables could be made by computer, at least 13 readability formulas (plus four variants of these formulas) could be automated. The three variables are: sentence-length in words; word-length in syllables; and words appearing on a prescribed list. These variables fall into two overlapping groupings. The first includes sentence-length and word-length, the input for the well-known Flesch Reading Ease formula; the second includes sentence-length and word-list appearance, the input for the popular Dale-Chall formula. The rationale for this approach was described in "Automation of Readability Measurement," a paper presented by George R. Klare as part of a symposium at the New England Educational Research Organization (NEERO) meeting (held at Boston College during June 5 - 6, 1969).

The first work was done on syllable (word-length) and word (sentence-length) counters, so that the Flesch and related formulas could be automated. The rationale for the syllable counter was described in a paper presented at the NEERO meeting entitled "Development of a Syllable Counter" (Paul P. Rowe). The first publication of the above work was Technical Report No. I.1.1: Automation of the Flesch "Reading Ease" Readability Formula, CAI

Version (FRECAI), January 20, 1969 (G. R. Klare and M. G. St. John). This work was expanded so that it eventually covered the Wheeler-Smith, Farr-Jenkins-Paterson, Gunning, and McElroy ("Fog Count") formulas as well. In addition, the following features were included to make the program as generally and widely useful to educators as possible: common hardware and software were used; a common programming language and logical program design were used; widely desired options were provided to handle various sample sizes and texts; and various related language counts were also provided. The program provides very highly accurate and reliable syllable (and overall) counts, and a procedure is being used to incorporate changes (when errors are found) to increase the accuracy of the syllable counter still further. This work was written up in a paper entitled: Automation of Flesch "Reading Ease" Readability Formula, with Various Options (G. R. Klare, P. P. Rowe, M. G. St. John, and L. M. Stolurow) published by the Reading Research Quarterly (Summer, 1969 issue, Volume IV, Number 4, pages 550-559).

The syllable counter will soon reach the very high state of accuracy required for use in the construction of instructional materials. This will, as indicated previously in the report, permit generative as well as analytic use of the program. For example, syllable-counted lists of words or passages can be presented to students as "word-attack" practice material for syllable counting (the work can be individualized by grading the words in terms of appropriate difficulty). Or, as another example, a sequence of uncounted, then counted, words can be presented in

sequence as a test of a student's skill.

The next work planned in this area (in addition to refinement of the accuracy of the syllable counter) is automation of a word-list look-up procedure. This will make possible the automation of the Dale-Chall, the Spache, and other formulas which use a word-list. Some work had been done previously in generating by computer a list of variations (i.e., common endings) of the root words in the Dale List. These will require checking and verifying, however.

(2) Cloze measurement. The "cloze" procedure is a relatively new measurement method which involves the deletion of every nth word of text (usually every 5th), with the student required to fill in the blanks. It is not possible to review the many uses of the method here. Briefly, however, it has been found useful primarily for measuring the difficulty of reading passages, or measuring an individual's reading skill. It is also highly flexible in that it can provide "n" different close-to-equivalent test versions of the same passage that can be used for pre- and post-testing, etc. That is, if every 5th word is to be deleted, one version can have words 1, 6, 11, etc., deleted while another can have words 2, 7, 12, etc. deleted.

A computer program has been developed which will delete every nth word, where "n" can range from 2 to 99. A related program deletes special characters and operations codes used in CAI programming so that the cloze program can be used with CAI materials. This work was described in a paper presented at the NEERO meeting entitled "CLIP - A Cloze Instructional Processor" (M. Gregory St. John). A more complete report will be prepared covering this work.

Future plans involve possible use of equivalent cloze forms as a method of evaluating amount learned from CAI lessons. A pre-test of a lesson, then instruction on the computer, and then a post-test of the lesson is currently being considered. Another possible analytic application involves the collection of dialect synonyms or alternative phrases through use of cloze versions of passages with students who use a particular dialect. Where a special expression is used in a dialect it should therefore be picked up.

The generative use of the cloze procedure is also being investigated. An experimental lesson has been developed by an instructor which uses student responses to word deletions as the basis for branching in the instructional program. This lesson has been found very useful with ABE students, and additional lessons of this type will be developed. As it becomes possible to describe the words an instructor deletes more precisely and objectively, an attempt will be made to integrate the necessary procedures to do this automatically into the cloze program.

(3) Speed of reading measurement. This type of measurement has not been required to date, and consequently no work has been done in this area.

(4) Acceptability (attitude) measurement. Three forms have been developed to measure student and teacher reaction to CAI materials. Two are student forms called "Scale of Attitudes Toward Computers for Boston Students," with both PRE and POST forms available. These are being given to each Learning Center student prior to first computer use (PRE), after each has completed his first lesson (POST), and again when each has completed his CAI training altogether, i. e., leaves the ABE portion of his training (POST).

The form for teachers is called the "Teacher Questionnaire for

CAI Unit Lesson. It is being filled out by the teacher or console assistant as each student completes each CAI lesson. Responses to the above forms will be analyzed, as indicated earlier, after the training period ends in July. Plans are being made to revise the forms as indicated by the first year's experience with them.

II. Word Changes. This heading is meant to include work on Project CREATES along the lines of word analyses. These are designed to lead to word changes (e.g., substitutions) which can increase instructional efficiency.

(1) General Inquirer and SOCCEP trials. The term "General Inquirer" refers to a content analysis system (or set of programs) of potential value in a text analysis and revision project such as that being carried out (see Stone, P. J., Dunphy, D. C., Smith, M. S., and Ogilvie, D. M., The General Inquirer: A Computer Approach to Content Analysis, Cambridge, Mass., M. I. T. Press, 1966). A particular set of programs in the system makes it possible to "disambiguate" a word, i.e., determine its meaning in a particular context. For example, the word "run" may mean something like "hurry", or "stream" (i.e., Bull Run), or "score" (in a game), among its various possible uses. By examining the words surrounding a given word, it is possible to set up rules based on the various patterns found so that it is possible to specify which meaning is intended. (An example of this process is presented in Technical Note No. 1: An Example of Proposed Text Analysis.) For later work in rewriting, this capability would be highly useful. To date, a number of conversations have been held on such procedures, and beginning plans laid. These involve eventual disambiguation rules for the kinds of words that would be used in an automated rewriting system.

Work has been limited in this area recently by the fact that other, higher priority tasks have been carried out, and also by the fact that the Inquirer II System has not been operating properly. (The Inquirer II is the version of the General Inquirer which will operate on the IBM System 360, a desirable system for the intended purposes.) An example of the kind of work planned with the General Inquirer (or Inquirer II) approach is given, as indicated earlier, in Technical Note No. 1: An Example of Proposed Text Analysis.

Related to certain disambiguation procedures that can be performed with General Inquirer programs is the use of programs which can provide key-word-in-context, or KWIC, listings. These are line-by-line printouts which show every occurrence of a given word within a text and the words which surround it. Such programs can be used to prepare concordances and indexes, code information for retrieval, etc. Two such programs have been run. The first is SOCCER (Smart's Own Concordance, Extremely Rapid),developed by Guy Hochgesang, and the second is BACON (Basic Concordance), developed by David Packard (and given the code name BACON for CAI Lab purposes).

(2) KWIC and frequency listings. Some limited work has been done with both the SOCCER and BACON programs along the lines of providing KWIC and frequency listings. The purpose of the KWIC listings is twofold. First, the analytic use of providing a means of disambiguating words has been planned. As indicated above, only some limited trial work has been done along these lines. Second, the generative use of providing new instructional materials has also been planned. For example, it is possible to use selected KWIC listings as a means of teaching students to learn different word meanings from context. Also, it is possible to teach students the

proper distinctions in usage of such words as "was" and "were" or, at a more complex level, "would" and "should", for example. Or, if the key word is itself deleted from KWIC lines, such listings can serve as tests of a student's knowledge of correct usage. Such applications of KWIC listings are shown in the diagram entitled: Examples of Some Current Program Output.

The SOCCER program, in addition to providing KWIC listings, also provides data on the frequency with which the words are used in the text being analyzed. This provision makes possible the use of the program for "vocabulary control" in the rewriting of passages for readers of limited ability.

(3) Sense rules. This term refers to the development of disambiguation rules to cover the various senses or meanings in which words occur in print. As indicated, work has been limited to date by other priorities and problems with the Inquirer system. As the need for sense rules becomes more critical later, further work will be devoted to this task.

(4) Word length and frequency changes. One of the goals of the revising and rewriting system that is planned under Project CREATES is the automation of certain word changes in text to increase the readability of the text. The first step to be undertaken, however, was a complete review of the literature to determine which variables in words were most likely to be effective in changing readability. This review indicated that the following were important:

(a) Word frequency, especially as related to reading speed and acceptability.

(b) Frequency of content words, particularly nouns.

(c) Concreteness (as opposed to abstractness) of words.

(d) Meaning-value (in terms of using words which have high association strength).

(e) Length (as measured in letters or syllables).

The review was completed and presented as a major part of Technical Report No. B.1.1: Words, Sentences, and Readability, May 5, 1969 (George R. Klare). This material became the basis for a subsequent report to be used as a part of the first version of a rewriting system. This is Technical Report No. B.2.1: A First Version of a Computer-Aided Revising, Editing, and Translating System (CARET I), June 15, 1969 (G. R. Klare, P. P. Rowe, M. G. St. John, and L. M. Stolurow).

III. Sentence Changes. This heading is intended to cover work on Project CREATES along the dimension of sentence analyses. Such work has been designed to lead to characterizations of sentences which contribute to their difficulty and ultimately to changes which can increase instructional efficiency.

(1) WISSYN program trials. The term "WISSYN" refers to a computer program developed at the University of Wisconsin to analyze syntax, i.e., to parse sentences. At the inception of the Project, it was felt that this program would provide the best available approach to sentence analysis. Descriptive material on the program was obtained, but further review of available materials by CAI Lab personnel and consultants suggested that other programs and approaches might be preferable. Consequently, a copy of the WISSYN program itself was never obtained.

Two other approaches to getting parsing information were investigated. First, the General Inquirer system's disambiguation procedures were designed to provide some information of this sort.

This made an additional program like WISSYN, for which a good deal of preparatory work would have been necessary before use, assume a low priority.

Second, interest grew during the course of the work in possible parsing information that could be derived from a dictionary on tape. Both the unabridged Random House and the Webster's Collegiate and Pocket dictionaries were investigated. Descriptive information on the Webster's dictionaries has been reviewed, and preliminary plans made to order the tapes for them if permission is granted. Many other possible uses of the dictionary are planned (e.g., definitions, spelling, etc.) besides the parsing information mentioned above.

(2) Form-class listings. Work in this area was to be dependent upon the parsing programs available for use. Very little direct work has therefore been done at this point. Future work will, of course, depend upon progress in the above category.

(3) Pronoun changes. A long-recognized source of difficulty in written material lies in the writer's use of pronouns. Among the problems requiring attention are: matching a pronoun to its proper referent: use of the "editorial we"; use of indefinite pronouns; etc. Pronouns may best be included under the category of "anaphora," or constructions (pro elements) referring to previous elements (antecedents) in a sentence. It is now clear that such constructions might therefore more properly have been included under the following major category ("IV. Other Changes"). Pronouns, and anaphora generally, were discussed briefly in Technical Report No. B.1.1: Words, Sentences, and Readability, May 5, 1969 (George R. Klare).

(4) Sentence length and complexity changes. As indicated for word changes (above), one of the goals of Project CREATES is the

automation of certain sentence changes in textual material to increase its readability. This is to be carried out by means of a revising and rewriting system under human control, but with the computer performing as many functions as possible. The first step to be undertaken was a complete review of the literature to determine which variables in sentences were likely to be most effective in altering readability. The review indicated that the following should be considered:

(a) Length, including both sentence length and independent clause length.

(b) Nominalizations versus active verb constructions (nominalizations are words or compounds used as nouns).

(c) Question versus statement constructions.

(d) Affirmative versus negative constructions (i.e., "not" or other negative words).

(e) Active versus passive constructions.

(f) Embedded versus non-embedded constructions (i.e., repetition of rules within a sentence, sometimes referred to as a recursive construction).

(g) Depth of words in sentences (i.e., the number of "commitments" each word incurs as part of a sentence, as based on the analysis developed by V. H. Yngre - see "A Model and Hypothesis for Language Structure," in Proceedings of the American Philosophical Society, 1960, 104, pp. 444-466.

Completed and presented as a major part of Technical Report No. B.1.1: Words, Sentences, and Readability, May 5, 1969, this research review became the basis for a subsequent report to be used as a part

of the first version of a rewriting system. This system is presented in Technical Report No. B.2.1: A First Version of A Computer-Aided Revising, Editing, and Translating System (CARET I), June 15, 1969.

IV. Other changes. This heading is a general one designed to include changes that do not fit easily into the word and sentence categories discussed above. The purpose is the same: to lead to knowledge of variables that affect text difficulty, and ultimately to make changes that can increase instructional efficiency.

(1) Nominalization changes. Studies have shown quite clearly that a nominalized construction is more difficult for a reader to comprehend than an active-verb construction. An example would be Noah's construction of the Ark versus Noah constructed the Ark (or better yet, Noah built the Ark). Nominalizations have certain characteristics which suggest grouping with Word Changes - section II - and certain other characteristics which suggest grouping with Sentence Changes - section III. For this reason they were originally included under Other Changes - section IV. During preparation of Technical Report B.1.1 it was decided that nominalizations could best be placed with Sentence Changes. Consequently, they were included under section III (above) and will not be further discussed here.

(2) Technical term definitions. Rewriting text to make it more readable involves, as one major element, substitutions for difficult words. Such changes can only be carried to a certain point, however. One limit recognized previously by the Project investigators involves technical terms. Since such terms have specialized, generally-accepted meanings, it appears better to teach their definitions than to make

substitutions for them. Another limit, which could be described more clearly after the Project began, is the "elemental" word (definitional primitive) for which no simpler substitute can be found. (Certain words can be replaced by common short phrases, but others cannot be handled in this fashion.) In such cases, once again, definitions are needed if the reader cannot understand the words.

A possible solution, it appears, is to provide a dictionary and associated look-up procedures for such words. This is one of the major reasons for the investigation of dictionaries available on tape (mentioned earlier under section III, item 1).

(3) Phrase and clause pattern difficulty. Readability research indicated early that the number of prepositional phrases in a standard-sized sample of text was one index of its probable difficulty to readers. Similarly, the ratio of complex and compound-complex to simple sentence constructions provided another index. It was felt that large samples of text might be analyzed by computer program to determine the difficulty of such patterns and establish rough norms as to when the various patterns might best be introduced into adult basic reading material. The review of work being done in transformational grammar suggests that such an approach might prove efficient and effective.

(4) Transformation difficulty. The advent of "transformational grammar" as a field of study was quickly followed by experimental studies of the relative difficulty of the various transformations. Studies have been reviewed and are presented briefly in the previous section (III, item 4). They are also described in a more complete fashion in Technical Report No. B.1.1: Words, Sentences, and Readability, May 5, 1969, and in Technical Report No. B.2.1: A First Version of a

Computer-Aided Revising, Editing, and Translating System (CARET I), June 15, 1969.

Another use being considered is an outgrowth of Technical Report No. B.1.1.: Since a computer-aided analysis has shown the increased use of optional transformations to be related to comprehension difficulty, such an analysis could be useful in several ways. Of greatest importance, perhaps, would be the establishment of rough norms for the presentation of such constructions as passives, negatives, interrogatives, etc. These could then be used in developing scaled reading material for adult basic learners. This should help to prevent frustration among readers yet, at the same time, provide material that is sufficiently challenging so that readers can increase their skill.

A. Instructional Inputs. This section covers the selection and preparation of a corpus, or body of reading material in punched-card form. This work has both practical and developmental importance. That is, a corpus can be used for the provision of instructional material in reading for adult basic learners in an immediately practical way. At the same time, it provides material that can be used in trial runs of various programs and procedures being developed. As it is refined over time, it should eventually provide a corpus of calibrated material of wide applicability in the field of adult education.

1. Grade level by interest corpus. This heading refers to a body of reading materials calibrated in the form of a 5 x 5 matrix, with interest category along one axis and reading grade level along the other. Materials have been selected from various sources to fit into this matrix; at present, approximately 65 lessons (reading stories)

totalling about 39,000 words have been punched. Other stories will be added as needed or desired.

The five categories of interest used in the above matrix were selected on the basis of studies of preferred adult reading interests, and include the following:

a. Health

b. Personality

c. Sports

d. Laws and Legislation

e. Racial and International Problems

Each of the above is purposely broad enough so that it can be broken into further categories as needed, on the basis of further research or application.

The five reading grade levels used are:

a. Below 4

b. 4 - 6

c. 7 - 9

d. 10 - 12

e. Over 12

Once again, these levels were based on rough categories frequently found in published adult basic and (at the higher levels) adult reading materials. They are wide enough in range so that each can be further sub-divided. At this point, the lessons have been assigned to the matrix on the basis of the rough gradings given by the publishers, and therefore no finer breakdown is justified. Once the readability programs being developed and refined under Project CREATES have been

applied to the materials finer breakdowns can be made, if needed.

A technical note on the preparation of this matrix has been prepared for office use only, but is included in the appendix for the interested reader. (No. 7: General Reading Material Matrix, 9/17/68, G. R. Klare). In addition, Technical Note No. 3: A Preliminary Study of the Reading Interests of OIC Learning Center Students, March 29, 1969 (G.R. Klare) has been prepared as a first study of the particular interests of Learning Center students. It is being used as a guide to priorities in the development of new CAI lessons for Learning Center use.

(2) Semi-technical corpus. Originally the Laboratory planned to develop a corpus of semi-technical material to parallel the above grade level by interest corpus. Three books or booklets were obtained from OIC (for the clerical and drafting areas), and one passage of approximately 777 words was punched. Further effort along this line was deferred, however, since it was felt that the first effort should go to the more general grade-level by interest corpus. Since the Project is an adult basic education endeavor, the more general materials were judged likely to be of wider applicability in later ABE use throughout the country.

It was felt, however, that work in this area should be resumed at some later time for two reasons:

(a) Pre-vocational materials (which is the meaning intended by "semi-technical" here) are potentially among the most interesting to adult learners. It might be possible to tap motivation with such materials in certain cases where the more general materials have failed. (These should not be considered as providing replacement

for all of the more general materials, however, since a reading lesson using, for example, the topic of "clerical filing" would probably be of interest only to those planning to enter or considering a clerical position.)

(b) If it is possible to prepare job-related materials at a low enough reading level to be learned by the better adult basic students, job training might begin earlier than it does at present. The extent to which this might be done is not clear at the moment, but certainly some improvement in readability is frequently possible. An illustration of this can be found in Technical Note No. 1: An Example of Proposed Text Analysis, October 1, 1968.

(3) General Inquirer corpus. In conjunction with work on the General Inquirer system, particularly that for disambiguation purposes, a corpus of over 500,000 words was prepared and made available for certain uses. Plans were made to use printouts from the General Inquirer corpus for disambiguation of words for word-change (substitution) purposes. Automated rules might thus be developed so that the computer could be used to disambiguate the 3,500 root words that lie between the 1,500th most frequent to the 5,000th most frequent in general counts of word usage. The General Inquirer project has been completing disambiguation rules for approximately the 1,500 most common root words, so these are potentially available. Plans to stop at the 5,000th most frequent word rather than go farther arose from a combination of circumstances:

(a) The most common 5,000 root words cover an estimated 97% of running text at typical adult reading levels. It seemed likely

that, if anything, the figure might be higher for adult materials of the sort to be used in the Project.

(b) Practical considerations suggested that to go much beyond 5,000 words in frequency would be inefficient for early work, since including 5,000 additional words would probably increase the percentage of text covered by only about two points or so, according to estimates. The results of early work could, of course, lead to revised estimates and further analyses where needed.

Some sample materials were obtained from the General Inquirer project staff and they were examined for possible use in disambiguation. It was decided at this point, for priority reasons, to switch attention to a rewriting approach which initially involved more human editorial emphasis than machine assistance. This led to the preparation of Technical Report B.2.1: A First Version of a Computer-Aided Revising, Editing, and Translating System (CARET I), which could not otherwise have been completed by the end of the first 9 1/2 months of work. This editorial approach thus became available for the second year of the Project, which would not have been possible if the approach involving more computer assignments had been given priority, since this latter is a longer-range task. Computer methods have a most valuable potential, however; thus, as soon as possible, attention will be shifted to the approach which involves the computer more fully. A brief picture of the more fully automated approach can be found in Technical Note No. 1: An Example of Proposed Text Analysis, October 1, 1968. Next steps planned will also be described in the

following section (B, item 2).

B. Rewriting systems (automation levels). This heading refers to the work done in the study and development of systems for re-writing instructional material to make it more readable. The first stage of the work has been completed, but much further work is needed and plans have been made to carry it out.

(1) Computer-assisted explanation. This is the name given by Edmund Berkeley to his set of computer programs designed to assist the writer (particularly the technical writer) in achieving clear explanation. The authors have noted elsewhere their indebtedness to Mr. Berkeley for ideas from his published work and for his personal help and encouragement in their present work.

Mr. Berkeley developed programs for counting syllables in words and words in sentences, and also for tagging certain categories of words for later editorial attention. In addition, he provided a number of ideas on how a writer or editor might make changes to improve readability - an art he himself has practiced very well, indeed. Since Mr. Berkeley's programs were originally written in ISP and designed to operate on a PDP-7 computer and later in machine language for a PDP-9 computer, it would have been very difficult for the authors to use them experimentally without considerable time and great expense. Mr. Berkeley thereupon donated time on his own PDP-9 computer, inviting the authors to use his computer and programs, with only previous notification of desire. He also served as a valuable critic of the ideas being developed by the authors. In sum, the authors borrowed time and ideas from him, and made good use of them in their own work.

As a parallel to the explorations with Mr. Berkeley's work, extensive background research was done along the lines of the kinds of language changes which should be made when improved readability is desired. This work culminated in Technical Report B.1.1: Words, Sentences, and Readability, May 5, 1969. Both of the above approaches are also discussed in the following item (B.2).

(2) Computer-assisted revision. The authors' first 9 1/2 months of work on a computer-assisted revision system culminated in Technical Report B.2.1: A First Version of a Computer-Aided Revising, Editing, and Translating System (CARET I), June 15, 1969. This report is in the form of a manual describing the two major parts of the CARET I system: the computer-aided format for rewriting, and the suggestions to the editor making changes to increase readability.

The following plans for the system have been made for the second year of the Project:

(a) Use of the system during the next year by a number of interested adult basic education workers for purposes of critical evaluation. A special evaluation form listing the features of the system has been developed for this purpose, and will be revised as needed. The reactions of such workers toward the current features and planned next features (plus their own suggestions) will be used in developing the second version of the system (CARET II).

(b) Use of the system by several curriculum specialists employed on the Project during the next year. They will also be asked to evaluate the present and proposed features of the system, and to supply their own suggestions. These will then also be used

in the development of the second version of the system (CARET II).

(c) The curriculum specialists will be using the system to rewrite instructional materials to improve their readability. Thus, instructional materials will be under preparation during the year for possible widespread use in adult basic education centers throughout the country. (This will be described more fully below.)

No attempt will be made to describe the system fully here, since a copy of the manual (Technical Report No. B.2.1) is included in the appendix of this report. Instead, a copy of that part of the evaluation form listing the features of the system is presented here for summary purposes in Figure 2.

Figure 2

Evaluation Form

CARET I System

	Rating				
System Features	Very Useful 1	2	No Opinion 3	4	Not Useful 5
1. Format					
a. Triple-spacing					
b. Sentence line-up					
c. Syllable count					
d. Word count					
e. Readability score					
f. Hardware and software usefulness					
2. Guidelines					
a. Considerations (e.g., motivation, etc.)					

Figure 2 (cont.)

System Features	Rating: Very Useful 1	2	No Opinion 3	4	Not Useful 5
b. Readability information (e.g., word data, etc.)					
3. Word changes					
a. Word-frequency familiarity					
b. Content-word frequency					
c. Word length					
d. Concreteness-abstractness					
e. Association-value					
4. Sentence changes					
a. Sentence length					
b. Nominalizations vs. active verbs					
c. Questions vs. statements					
d. Affirmations vs. negatives					
e. Actives vs. passives					
f. Embedding					
g. Word depth					
5. Future possibilities					
a. On-line revising					
b. On-line readability					
c. List of alternative words					
d. Definitions					
e. Synonym substitution					
f. Tags for sentence changes					
g. Possible sentence changes					

C. Instructional outputs (applications). This heading is intended to refer to the use of the various programs and systems developed during the Project work to date in helping to provide instructional materials geared to the needs of adult basic education students. As indicated earlier, these various programs can have both analytic (item 1 below) and generative (item 2 below) functions, and are to be part of a more extensive set of programs to be called PAGES (for Psychological-Analytic-Generative Educational System). Preliminary work has actually begun under each of the items below, but major emphasis will be placed in this area during the second year of the Project.

(1) Evaluation of existing adult materials. Two chief analytic methods of evaluating existing adult materials were developed and plans were made to modify a third during the period to date: the readability formula program for applying the Flesch, Gunning, Farr-Jenkins-Paterson, McElroy, and Wheeler-Smith formulas; the cloze procedure; and the concordance procedure. Plans have been made to apply the readability program to:

(a) All CAI lessons now available on the system at Harvard University. Such an application has already been made, but with an earlier version of the program that was not as accurate as possible. Since the program is continually being increased in accuracy as sources of error are discovered, further applications are planned.

(b) All reading materials in the grade level by interest corpus. As previously mentioned, the only readability scores available on these materials are publishers' estimates, so further analyses are called for.

(c) As possible instructional materials are selected or developed by Boston University for adult basic use, some of these, at least, will be analyzed for readability. Tentative plans have been made to have OIC students who are in keypunch training punch the stories as part of their instructional work. If this arrangement were to work out satisfactorily, far more material could be handled than would otherwise be the case.

Before the cloze technique can be widely applied to adult basic materials it will be necessary to carry out certain trial work with it. It is not known, for example, what deletion pattern will prove most satisfactory. It seems likely that deletion of every fifth word will provide material too difficult and frustrating for adult basic students. Whether a pattern of, for example, 7 or 9 or 11 would be better from the point of view of difficulty, yet also be satisfactory psychometrically, remains to be determined. Fortunately, the cloze program that has been developed provides for any desired pattern of deletion.

Once the best cloze pattern has been determined, it will be possible to apply the cloze program to any of the materials mentioned above. Because there are "n" different versions of an nth deletion pattern, and because the different versions are roughly equivalent in difficulty, it is possible to supply pre-test and post-test versions or other kinds of roughly equivalent forms with relatively little difficulty.

The concordance, or key-word-in-context procedure is needed for disambiguation essential to later word changes. An additional analytic function it can perform, however, is the presentation of

data that can be valuable to the writer of instructional materials. For example, a writer who wishes to control the vocabulary of a particular lesson he is preparing or examining, or that of a prepared lesson he wishes to try in class,can readily get complete data on all words used, how often they are used, where they are used, what meanings are being taught, etc. Thus the two concordance programs being evaluated can provide a tool of considerable potential to writers.

(2) New instructional techniques. This heading may be slightly misleading, since it actually refers as much to new educational materials as it does to new techniques. The intention here is to indicate that, wherever possible, the analytic programs being developed are intended to have the potential of generating new instructional materials as well. The purpose is to make it easy for an instructor to obtain special materials geared to his students' special needs and abilities. Plans include the following:

(a) Use of the readability program to provide practice materials for syllable counting. Since a syllable counter has been developed, it becomes possible to have a student practice on the words on one part of a computer printout, then check himself on a syllable-counted part that follows below. If more or less difficult words are needed, proper movement up or down the grade-level portion of the reading material matrix should meet such needs.

(b) The cloze program described above offers perhaps its greatest generative potential as an aid in the development of CAI lessons. It is possible to determine branching instructions on the basis of a student's answers to the deleted words in a lesson.

To the extent that the cloze program can be modified to delete those words an instructor might wish to have removed, it can provide at least a preliminary version of a lesson for the instructor's further modification. One step that might possibly be taken, for example, is the coupling of parsing information with the cloze program, so that only certain kinds of words are deleted. Another step is the storing of the words when deleted, so that they can be supplied as instruction when needed.

(c) The concordance program offers its greatest potential in helping the reader to learn proper usage or meaning from a word's context. By providing printouts of words in their linguistic environment (sometimes called key-word-in-context, or KWIC, printouts), it becomes possible (as mentioned earlier) to show a beginning student where "was" should be used rather than "were," or the way in which "run" can refer to a kind of movement, a score, or a brook, for example. The two concordance programs mentioned above (SOCCER and BACON) are to be modified, as possible and as needed, to provide printouts to meet the above kinds of applications.

(3) Rewritten materials for training use. As indicated above, plans have been made to begin the rewriting of instructional materials. This work should provide not only new materials for teaching uses but also a basis for evaluation of the CARET I rewriting system. This work will constitute one of the major emphases of the Project efforts during the coming year.

Sub-Project CAIBLS

As indicated earlier, this part of the Project CREATES report is not complete at this time. Because of the OIC Learning Center

schedule, training of the first group of ABE students will not have been completed until early July, and time will not permit data analyses to be included here. These analyses will be added at a later date, and will consist largely of the kinds of information to be specified at certain points below.

I. Materials development stage. This heading is meant to refer to preparation of instructional and analytic materials, and not to refer to systems development. Furthermore, it is a continuing stage, not one that will be completed at any given point. As information regarding the needs and interests of adult basic students is gathered and revised, the materials themselves are, and will be, revised accordingly.

(1) Preparation of instructional materials. This includes the preparation of the following kinds of materials.

(a) Lessons (in story format). A number of lessons have been prepared under the general heading of ALAN (Adult LANguage) CAI lessons. These lessons follow a format which, briefly, involves the presentation of text and then asks questions of the student. On the basis of his responses the student is branched to appropriate remedial material when needed. Records are kept of individual student responses for later use in instruction of the student or revision of the lesson.

Figure 3

Lessons on the CAI System

About Laws	About Environment
Law about False Advertising	Take Shelter
Social Security	Satellites
It Pays to Talk to Others	

Figure 3 (cont.)

About Jobs

Occupations I
Barbering
Calling about Work
Beauty Shop Operator
Employment Agencies
Hotel Work
Wanted: Skilled Workmen
Filling out an Application

About Health

Asthma
Food and Your Weight
Cancer and My Best Friend
You Can Help
A Fine Fix

About Consumer Education

Pasteurization of Milk

Some Materials Available

About Laws

Moonshine
Fishing Laws III
Child Labor Law

About Environment

Oil
Aluminum
Home Improvement

About Jobs

The Interview
Job Training
Shipping and Receiving Clerks
Church Ministry
Printing
Part-Time Sales

About Health

Medecine Can Be Harmful
Good Posture
The Flu
Smoking
Health and Pregnancy
Heart Attack
Mentally Retarded
Headache
Some Facts about Suicide
Just in Time

About Consumer Education

Interest II
Interest III
Double Talk
Mail Fraud

A new category of CAI lessons has been prepared called ALAN01. The first example is on the topic "Smoking" and is in an experimental format different from the ALAN series in that blanks are used to elicit responses instead of questions (as in ALAN). A major advantage of this type of format is that it opens lesson development to partial automation through modification of the cloze program mentioned earlier. A complete description of the ALAN01 format is

presented as the "Explanatory Note" in Progress Report No. 8, June 1, 1969.

(b) Reading and language skill programs. This heading refers to special skills that play a part in the overall reading process. Programs for the development of such skills had been planned for inclusion in the CAI Laboratory's offerings, but work on them received special impetus from the OIC Learning Center's request for such material. The fact that many of the Learning Center students had much more limited reading and language skills than was expected was the primary reason for such a request.

Special needs were expressed for a spelling program and a program teaching phonic skills. Programs for both skills are currently under development and should be available for use early during the second year of the Project. Both are being developed carefully so that they can be as flexible and broadly useful as possible. Other special skills programs are being considered, using a set of 10 teacher-made practice exercises in various areas selected so as to be of maximum usefulness in the Learning Center and for computer presentation. In addition, a special consultant in reading on the CAI Lab staff held several conferences with the Learning Center staff to aid in instruction of the students with severely limited language skills.

(c) Arithmetic programs. CAIBLS stands for Computer-Aided Instruction in Basic Language Skills. This suggests the actual emphasis of the work, which is instruction in basic language skills. Early in the Project, however, the Learning Center staff indicated that materials for teaching arithmetic skills would be

highly desirable, especially for students with severely limited language skills.

Consequently, two CAI Laboratory programs were made available, one on basic arithmetic operations (URAE02) and one on fractions (HALF). In addition, a specialist in the teaching of arithmetic held two conferences with the Learning Center staff. During the summer of 1969 plans are being made by the specialist to develop additional arithmetic materials. One of the unique processes to be studied is the generation of arithmetic problems through the use of a random-number generation program for the computer. Such a process would follow along the lines of the previously mentioned emphasis on generative programs, and would provide highly individualized instruction in the sense that no student need have exactly the same practice material as any other.

Other materials are also being considered, such as word problems making use of simple, everyday enumerating, counting, and computing situations. In addition further problems for both URAE02 and HALF are being readied for use.

(d) Tests. Just as instructional programs can be presented by computer, the computer's flexibility also permits the presentation of educational and psychological tests. Both visual and auditory presentation can be used, and detailed records automatically kept.

To date the following programs have been developed for computer presentation: The Porter (Job Corps) Reading Test; the California Achievement Test; and the Barnell-Loft series, as modified and adopted for CAI. Both fill-in and multiple choice versions are available. The Porter test (RP test) was found to be of limited utility because of the severely limited skills of the

students. It proved rather steeply graded in difficulty, and therefore quite frustrating to many. Furthermore, it appeared threatening to students who had, in most cases, already been defeated by the educational system and its ever-present tests. (However, it was found that the test could be of potential value with somewhat more skilled readers and will be retained for use with them.)

For the above reasons, it was decided that the Barnell-Loft series should prove more valuable, and it will be tried out during the second year of the Project. Furthermore, it was decided that tests for ABE students should generally be less steeply graded and game-like rather than test-like in nature, as the Barnell-Loft series is.

(e) Objectives development and review. Since the development of instructional objectives for ABE lessons was part of Boston University's contribution to the overall multi-agency project, the CAI Laboratory did not become involved in it. Laboratory members did, however, review the objectives developed by Boston University and attend the meetings at which they were discussed.

(f) Development of measurement devices. Four kinds of special measurement devices were developed, and three were applied during the first 9 1/2 months of the Project.

(1) The first kind of measurement device was an attitude form called "Scale of Attitudes Toward Computers for Boston Students," with a PRE form for pre-computer use and a POST form for post-computer use. (These are included in the Appendix.) Plans are underway to revise these forms for use during the second year of the Project.

(2) The second kind of device was an evaluation questionnaire for teachers who supervise CAI lessons, called "Teacher Questionnaire for CAI Unit Lesson." It was designed to be used by teachers during and just after each CAI lesson they supervise for each student.

(3) The third kind of device was a set of sheets describing titles and content of a set of adult lessons. They were given to Learning Center students, with necessary preliminary instructions, as a way of determining their reading preferences. The results were presented as Technical Note No. 3: A Preliminary Study of the Reading Interests of OIC Learning Center Students, March 29, 1969. These results are now being used in determining priorities for development of CAI materials.

(4) The fourth kind of measurement device that was developed in program form was the cloze procedure program. It has been described earlier in this report, so the description will not be repeated here. Plans are underway to use cloze tests both before (PRE) and after (POST) use of CAI lessons during the next year, to measure amount learned. Before this can take place, however, some study will be necessary to determine the cloze pattern that will be most effective with ABE students.

(2) Preparation of analytic materials (programs). This refers to the programs described under the heading Sub-Project TREE, particularly the readability, cloze, and concordance (KWIC) programs.

They will not, therefore, be described again here. It should simply be noted that these programs are intended to be of special value in the evaluation of instructional materials.

One kind of existing adult material that requires evaluation is that which has been prepared by our Laboratory, or by others in CAI form. This involves the addition of CAI operations codes and special characters for computer use. If codes and characters were counted as regular English text characters by our analytic programs, they would tend to distort many analyses.

It therefore became important to develop a special program to remove such codes and characters and "re-pack" the text so the analytic programs could be used to accurately represent the text. Such a program has been developed for CAILAN, the Harvard CAI Lab's interactive instructional programming language. It therefore opens all CAI programs on the Harvard system to the analytic procedures described in this report. (The program could, of course, be developed or modified for other CAI languages if the need arose.)

II. Student acceptance stage. This heading is meant to refer to the study of the acceptability of CAI materials to the students who use them.

(1) Measurement of acceptability to student users. The PRE form of the "Scale of Attitudes Toward Computers for Boston Students" was given to students prior to any terminal experience (the console assistant failed to give the form prior to the first experience of some of the early students, so these data can not be included). The POST form was given after the completion of the first CAI lesson, and then again at the end of the ABE training

period.

As indicated earlier, the results could not be analyzed at the time of this report, but will be provided in a later report.

(2) Measurement of instructional efficiency and value to student users. For the first training period for ABE students at the Learning Center, the measurement of instructional efficiency will be based on student responses and errors to the instructional programs (lessons). A summary of these data will be presented in a later report, and supplemented with subjective comments provided by ABE students on the value of the lessons to them. For the second year of the Project, it is hoped that such data will be further supplemented by PRE and POST cloze test scores on the passages. As indicated, a preliminary study of cloze will be necessary first, however.

III. Staff acceptance stage. This heading refers to the acceptability of the materials developed to the Learning Center teachers who are concerned with their use.

(1) Evaluation by teachers as to instructional efficiency, appropriateness, etc. As indicated previously, the "Teacher Questionnaire for CAI Unit Lesson" was completed by the person assisting the students (usually the "console assistant," who was a teacher's aide) after each lesson was completed by each student. These data will be supplemented later, when available, by subjective comments from the Learning Center staff.

(2) Determination by teachers and teacher-trainers of how and when materials can best be used. Discussions were held throughout the first training period for ABE students from January to July.

Various suggestions for use of the materials were made, and these eventuated in several kinds of action.

(a) Accelerated development of special-skill programs, of the type described above (spelling and phonics materials).

(b) Presentation of arithmetic programs, of the type described above (basic operations and fractions materials).

(c) A statement of the kind of CAI materials and their characteristics likely to be most useful and effective with ABE students. This statement was presented as the "Explanatory Note" included in Progress Report No. 7, May 1, 1969. It is hoped that continued discussion will lead to further additions to, and modifications of, this statement.

IV. Curriculum stage. This heading refers to the coordination of the CAI-developed materials with other materials that are to be, or are being, used in an ABE program that is actively in progress.

(1) Relation of materials to an organization's on-going program. As indicated above, special attempts have been made to fit the CAI materials being developed into the OIC Learning Center's program needs. The following approaches (most of which have been mentioned above) have been, and are being, used.

(a) Development of reading lessons according to the interests and ability levels of ABE students. The use of interest study data and the application of readability formulas take the guesswork out of the process and make it as objective, scientific, and widely applicable in other Learning Centers as possible (see section V below for further details).

(b) Provision of special skills program (e.g., spelling

and phonics) in the language and reading area to meet the reading needs of readers with very low-level skills. These programs are not being rushed to completion, but are being developed with sufficient care that they can be applied in other ABE programs as well as the OIC Learning Center's work (see section V below).

(c) Provision of special arithmetic programs to meet OIC Learning Center needs, especially for ABE students with very limited reading skills. Once again, special attention is being given to flexible arrangements in programming that will provide for more general applicability. Everyday situations and needs are being used as the lesson framework in order to provide increased student motivation and interest.

(d) Modification of the educational and psychological tests being developed for use with ABE students. Tests with a game-like rather than test-like appearance are to be used, especially with the lowest-skill students who are most likely to be frustrated by the test situation. Higher-skill students, according to the Learning Center staff, expect and accept more traditional tests, and such tests can therefore be used with them.

(2) Determination of how and when materials can best be used. Two general approaches have been followed in this area.

(a) Communication with OIC Learning Center staff. As indicated previously, conferences for communication purposes have taken place frequently (every two weeks) most of the year. In addition, comments have been solicited from the staff on the information and plans provided in the "Explanatory Note" section of the Progress Reports. Also, regular meetings have been held with the

console assistant at the Learning Center, usually at least once a week. Finally, visits have been made by various CAI Lab staff members to the Learning Center. These avenues of communication have led to some changes that are not recorded explicitly; others, of a more general and widely applicable nature, have been recorded in "Explanatory Notes" and this report itself.

(b) A questionnaire has been developed and will be tried out shortly on a trial basis at a regional meeting of ABE personnel. It will provide objective data on the kinds of materials, equipment, and procedures being used in ABE and, more importantly, those felt to be still needed by ABE personnel at the OIC Learning Center and elsewhere. Emphasis will be upon "how" and "when." As the questionnaire is refined, it will be possible to use it in a national survey if the opportunity can be provided. This will be an important objective supplement to the information gathered by CAI Lab personnel in attendance at local, regional, and national ABE meetings.

V. Generalization stage. This heading refers to the application, in a more general setting, of the materials developed, and being developed, at the CAI Lab.

(1) Application to similar organizations. Arrangements are now being completed for use of ABE materials in the Boston Public School Adult Basic Education's WIN (Work Incentive) program. Discussions have also begun with other ABE training program directors in the Appalachia area and in two large Eastern cities - also, CAI Lab staff recently collaborated on the preparation of an ABE Newsletter devoted to possible use of CAI programs for ABE training in industry. The interest generated at a recent national

meeting suggested that other ABE training programs might well become interested in CAI when knowledge of the materials and procedures developed becomes more widespread. With this kind of application in mind, members of the CAI Lab are planning further written statements and attendance at meetings to let other ABE projects know of the work being done.

(2) Use in hard copy form. As indicated earlier, one focus of attention for the work of this coming year will be the development of ABE training lessons in hard copy form. This is becoming possible through the development of the analytic-generative programs described in the TREE section of this report. Present plans envision the following rough steps in hard-copy lesson development.

(a) Selection of topics on the basis of information on ABE students' interest, as recently studied (see Technical Note No. 3: A Preliminary Study of the Reading Interests of OIC Learning Center Students, March 29, 1969, for a first study of ABE interests).

(b) Selection of potential lesson materials in the above topic areas. These may potentially come from certain commercially published adult materials, from U. S. Government publications, from general publications (in the mass media), from the CAI Lab's own materials, etc. It is essential that the materials selected be seen as relevant by ABE students. An evaluation of this should be greatly aided by the questionnaire developed to review ABE materials, methods, and procedures now being used in the developmental phase (see the description in the previous section, IV, item 2).

(c) Readability analysis of the potential materials. This will be done using the CAI Lab's readability program developed for this Project (see Automation of the Flesch "Reading Ease" Readability Formula, with Various Options, published in the Reading Research Quarterly).

If materials are at a proper reading level for ABE use, it will be possible to skip the next step. If not, or where additional materials are needed for the overall program, the next step will be used. If the first revision proves not to have reached a desired reading level, readability formula program application will indicate further revision is called for.

(d) Rewriting of potential ABE materials. Where rewriting is needed, the process will involve use of the system developed for this Project (see Technical Note No. B.2.1: A First Version of a Computer-Aided Revising, Editing, and Translating System (CARET I), June 15, 1969).

It is hoped that several versions of a lesson at different reading levels can be developed, so that ABE instructors can assign material on the same topic to all students in a class, for later discussion purposes.

(e) Once the material is found to be at the desired reading levels, the cloze program now being developed (and complete in some versions) will be applied. Several major uses will be made of it, according to present plans.

(1). Several cloze versions of the passage will be developed so that one or more will be available for both PRE and POST use in measuring amount learned.

(2). A CAI version of the lesson will be developed,

using student responses to blanks for instructional branches. This will be used for instruction and for refinement at CAI terminals.

(3). From the above, hard copy versions of the lesson will be developed in which the blanks and instructional branches can be provided in a kind of programmed instruction form.

(f) A set of hard copy materials will be developed consisting of:

(1). Readability data on the different versions of the passage so the instructor can select material at the proper reading level for his different students.

(2). PRE and POST cloze tests on the different versions, to measure amount learned.

(3). The instructional lessons themselves in hard copy, programmed instruction form.

(4). Vocabulary control data (in some cases), so that an instructor can teach special words or meanings in the passages if he desires. This data would come largely from the concordance program (BACON) being modified for use in the Project. This program has been described earlier in this report.

Though the above plans are still tentative, and subject to additions and modifications, they show how the Harvard CAI Lab's Project CREATES provides a coordinated attack on the problem of instructional material for ABE programs, in hard copy as well as CAI form. The authors hope that the plans also show how the various strands of the project mentioned in this report, though

seemingly unrelated in some cases at least, are being woven into an integrated approach to the preparation of improved adult basic education materials. Increased instructional efficiency can thereby be gained in a larger variety of programs. Also this approach provides a convenient and evolutionary approach to the management of individualized instruction. Thus the developmental plan of Project CREATES includes both educational accountability and coordinated management of the instructional programs of individual students.

This copy of the Annual Report does not have the appendix referred to on page 6 (This appendix contains most of the publications, technical reports, and other notes listed on pages 6 - 9.). Unfortunately, because of reproduction and mailing costs, the appendix was included in only a limited edition of the Annual Report. However, a complete report has been sent to the Educational Research Information Center (ERIC), and may be requested from them by writing to:

ERIC Clearinghouse on Adult Education
Syracuse University
107 Roney Lane
Syracuse, New York 13210.